Wang, H., & Yu, G. (2021). Test-Takers' Cognitive Processes during a Listen-To-Summarize Cloze Task. *International Journal of Listening*, *35*(1), 1-28. https://doi.org/10.1080/10904018.2018.1473040

Test-takers' Cognitive Processes during a Listen-to-Summarize Cloze Task

Authors: Haiping Wang & Guoxing Yu

1. Haiping Wang

Affiliation: East China University of Political Science and Law, School of Foreign Studies

Email: cathy_pingcn@163.com

2. Guoxing Yu

Affiliation: University of Bristol, School of Education

Email: guoxing.yu@bristol.ac.uk

Test-takers' Cognitive Processes during a Listen-to-Summarize Cloze Task

Abstract

Listen-to-summarize cloze (LSC) tasks are not rare in high-stakes language tests; however, we know little about what test-takers do during the tasks. In our study, we examined sixteen students' think-aloud protocols (TAPs) which were recorded while they were completing two LSC tasks. The analysis of the sixteen TAPs indicated iterative cognitive processes of LSC task completion. Simultaneously, the different gap types in the LSC tasks required different levels of cognitive process. We also discussed the possible cognitive depth to which LSC tasks can target and the challenges in designing appropriate gap types to assess different levels of cognitive processes.

Keywords: listen-to-summarize cloze (LSC), cognitive processes, think aloud protocol (TAP), academic lectures, listening comprehension

Test-takers' Cognitive Processes during a Listen-to-Summarize Cloze Task

Academic lectures as a principal discourse in academic settings (Benson, 1994; Lee, 2009; Richards, 1983; Waggoner, 1984) contain complex discourse markers and structures, and can pose a significant challenge for second-language listeners (Buck, 2001, p. 43). Various studies have demonstrated that second language learners with adequate language proficiency may find it difficult to comprehend an academic lecture (Hyon, 1997; Young, 1994). Students' academic listening competence has been increasingly recognized as an important pre-requisite for academic success (Flowerdew, 1994; Lynch, 2006). In some high-stakes language tests, such as International English Language Testing System (IELTS), Test for English Majors Band 8 (TEM-8), Pearson Test of English Academic (PTE Academic), etc., academic lectures are used as input or source in listening comprehension tests, with the aim to create a closer relationship between test tasks and target language use. In the listen-to-summarize cloze (LSC) tasks in IELTS and TEM-8, test-takers are required to listen to short academic lectures and fill in a few gaps of different levels of details and summarization of source lectures, as a way of assessing their academic listening competence. Compared to read-to-summarize tasks (for a review see Yu, 2005; Cumming et al., 2016; Taylor, 2013), listen-to-summarize tasks (e.g. Kintsch & Kozminsky, 1977; Kintsch & van Dijk, 1978) in the field of language testing are much under-researched. We know very little about what LSC tasks measure and what test-takers do during the tasks. In this paper, we report the cognitive processes of sixteen students completing the LSC tasks of TEM-8.

Literature Review

Listening comprehension involves a series of cognitive processes, incorporating both "bottom-up" and "top-down" processes (Anderson, 1995; Carroll, 1972; Clark & Clark, 1977;

Hansen & Jensen, 1994; Rost, 1994). "Bottom-up" and "top-down" processes might be employed simultaneously to constitute interactive processing (Goh, 2005, 2010). Field (2013, pp. 95-96) made a list of cognitive processes involved in listening comprehension, such as input decoding (acoustic cues transformed into groups of syllables), lexical search (best word-level matches identified for what has been heard), parsing (lexical sense specified in the co-text and a syntactic pattern imposed), meaning construction (world knowledge and inference employed to add to the bare meaning of the message), and discourse construction (relevance of the new information with the old and integration of the new information into a representation of the larger listening event). Taylor & Geranpayeh (2011, p. 96) presented a similar framework for listening test validation, which comprises goal setting, decoding acoustic/visual input, syntactic parsing, establishing propositional meaning, inferencing, building a mental model, creating a text-level representation and monitoring comprehension. In real listening comprehension tests, cloze tasks are not the test-takers' own words. They need to carefully read the tasks before, during and after the playing of the recording. Decoding process therefore includes decoding both acoustic and visual input (the task itself or notes of the recording). The task divides the test-takers' attention and becomes cognitively demanding (Pashler & Johnston, 1998). Moreover, researchers have stressed the importance of higher-level cognitive processes for successful listening comprehension (Anderson & Lynch, 1988; Brown & Yule, 1983; Buck, 1991; Rost, 1990; Taylor & Geranpayeh, 2011), yet little is known of what cognitive processes a listener would be engaged in when listening to an academic lecture (Field, 2011), and little attention is paid to listening of spoken academic genres (Flowerdew, 1994; Kiely, 2004; Lynch, 2011; Rowley-Jolivet, 2002). Field (2013, pp. 100-103) specified the higher-level cognitive processes in listening comprehension, i.e. meaning building and discourse construction. In order to

construct a holistic comprehension of a spoken discourse, four cognitive processes should be applied in discourse construction (Field, 2013, pp. 102-103), (1) listeners' selection of a new piece of information to the discourse as a whole; (2) integration of the new item of meaning into the developing of discourse representation; (3) self-monitoring of a new piece of information to ensure its consistency, and finally (4) the hierarchical structure building of what has been said. Field's model emphasizes the cognitive processes typical of comprehension of large stretches of discourse and these processes could also be essential for academic lecture comprehension (Field, 2009, 2011). Take selection as an example, listeners have to constantly select information for processing, due to their limited cognitive capacity. Consequently, they might fail to pay consistent attention to a speaker who is taking a long turn of speech, as in an academic lecture or a conferment talk (Brown, 2008, p. 11). According to Field (2013), discourse representation is dynamic and evolving while new input is being integrated; processes such as selection, integration and self-monitoring prepare for the last step – structure building which is rather a conclusive step that defines the quality of discourse construction and comprehension.

The challenge of developing an academic listening test is to ensure that "the cognitive processing activated in the test-taker by a test task should correspond as closely as possible to that in the real academic listening context" (Taylor & Geranpayeh, 2011, p. 96). As academic listening involves higher cognition (Hulstijn, 2010), academic listening tests should engage test-takers' higher-level cognitive processing (Taylor & Geranpayeh, 2011). Summarization tasks can serve that purpose as they are considered to be able to elicit higher-level cognitive processing (Yu, 2005; Taylor, 2013). However, Buck (2001) argued that the forms of summarization tasks (e.g. whether involving multiple-choice questions or gap-filling tasks) and the assessment focus of the gaps in the LSC tasks could easily alter the intended construct of the LSC tasks.

Furthermore, as we argued earlier (Yu, 2013), summarization is a hugely complex mental representation of source texts and it can serve multiple purposes which define the mental representation of the source texts to a large extent.

In terms of using summary-cloze tasks to assess reading comprehension, Taylor (2013, p. 79) has demonstrated that the summary-cloze has the advantage of assessing reading comprehension, because test-takers need the knowledge to correctly fill the gaps by reading and understanding the source text. Cloze tests are not limited to "local" comprehension only; they can test inferences, unlimited in scope (Gellert & Elbro, 2013). Comprehension-demanding gaps that target information integration ability across sentences should be designed to require readers to summarize information from more than one sentence (Greene, 2001). Storey (1997) provided convincing evidence of deep processing being generated by summary cloze tasks. However, Coniam (1993) argued that test-takers might fill in certain gaps without referring to the source text, which could raise questions on the construct validity of summary-cloze tasks. Furthermore, Courchêne & Bayliss, (1995) have shown that the length of source texts for summary-cloze tasks could further complicate the design of summary cloze tasks.

Yet only a few studies examined cloze tasks as a measure of listening comprehension (e.g. Buck, 1992; Lewkowicz, 1992). Lewkowicz (1992, pp. 18-19) found listening summary cloze tests involved two stages that demanded good note-taking skills and global understanding of the discourse, with "the added advantage of being a reliable testing instrument". Buck (2001, p. 73) further argued that the cloze task based on the summary of source text(s) could test: general linguistic knowledge, discourse knowledge and even inferencing, depending on the nature of the information to be inserted. Field (2009) studied the cognitive validity of lecture-based test items in the IELTS Listening paper (including summary-cloze tasks), in which he

argued that much of the cognitive processing in the test condition was shallow and at a "very local level", and that the LSC task in IELTS test was cognitively more demanding than note-taking in authentic lecture listening. In LSC tasks, if test-takers only need to decode local information to fill in the gaps, this kind of gaps usually elicits details directly from the recording. Other gap types (inference or synthesis gaps) in LSC tasks, furthermore, tend to target indirect inferences or even summary of the correlative recording and test-takers need to construct meaning to fill in the gaps. In reality, the sentence frame of the tasks may well paraphrase information from the recording, but the words to be inserted into the gaps are often taken verbatim from the recording instead of summarizing a larger unit than a lexical chunk (Field, 2013, p. 131).

Nevertheless, the nature of test-takers' cognitive processing (especially the higher-level cognitive processing) involved in LSC tasks with academic lectures as input/source remains largely unexplored, partly because of "the inaccessible nature of the skill" (Field, 2011, p. 102). It is imperative to explore these cognitive processes and their interactions by using in-depth qualitative research methods in order to better understand how test-takers achieve successful comprehension (Vandergrift, 2007, p. 206). The study reported in this article addresses these validity issues by analyzing the think-aloud protocols (TAPs) of sixteen participants while responding to the LSC task of TEM-8.

Method

Research Questions

1. What are the main cognitive processes test-takers undergo while completing the LSC task after listening to an academic lecture?
2. To what extent do gap types of the LSC task affect test-takers' cognitive processes?

3. To what extent does the LSC task target test-takers' higher-level cognitive processes?

Participants

Sixteen students from a key university in Shanghai (half in their junior/third year, and the other half in their senior/fourth year) participated in this research as volunteers. All were Mandarin speakers. Table 1 presents the participants' demographic information and the scores they achieved in the LSC task of TEM-8. The seniors took the LSC task of the 2010 version (see Appendix A) and the juniors took the LSC task of the 2013 version (see Appendix B). The LSC tasks used in the project are actual TEM-8 listening papers and they are assumed to be of equivalent difficulty based on expert judgements of the testing committee. We employed the two versions in order to cover all the gap types of the TEM-8 LSC task.

Insert Table 1

All of the participants were required to pass TEM-8 in their fourth academic year in the university. They were of similar age, ranging from 20 to 22 (mean=21.38). On average, they had learned English for 12.38 years, ranging from 9 to 15 years. It is also noticeable that the cohort was not balanced in terms of gender. Females (n=12) were dominant in number, compared with the number of males, only four, which largely reflects the overall gender ratio of the students in the university. The average score of the juniors group (2013 version) is 3.938 and the seniors group (2010 version) 3.688. The mean difference is -.25. According to the independent samples t-test results, there's no significant difference across the two groups' task performances in terms of their means, because the t-value is -.343 with the significance level of .736. In other words, the participants in two different year groups have approximately the same lecture comprehension ability level on average.

Research Instruments

TEM-8 is a large-scale, nationwide test developed and administered by the National Advisory Committee for Foreign Language Teaching (NACFLT) on behalf of the Higher Education Department, Ministry of Education, P. R. China, with the purpose to assess the English proficiency at the end of the advanced stage (the third and fourth year) of Chinese university students majoring in English Language and Literature (Jin & Fan, 2011). TEM-8 LSC is the first section of the TEM-8 listening paper that aims at testing students' listening skills to acquire and understand not only general information but also academic information concerning literature, linguistics, culture, etc. (NACFLT, 2004). It is administered in two stages: in the first stage, test-takers need to listen to a mini-lecture of around 1000 words for approximately 8 minutes while taking notes. In the second stage, they need to complete a cloze task (ten gaps) on an answer sheet and write no more than three words for each gap with reference to their notes, all within ten minutes. The response format of the TEM-8 LSC task is a summary cloze task in an outlined form (see Appendix A and B).

The TEM-8 LSC task is composed of four types of gaps based on expert judgements (NACFLT, 2004):

- gaps that target key points of the mini-lecture;
- gaps that target supporting details;
- gaps that target inferences;
- gaps that target syntheses of a group of sentences.

The four gap types target different sub-constructs of listening comprehension (see also Buck, 2001, p. 73). As shown in Table 2, inference gaps and synthesis gaps particularly aim at higher-level cognitive processes (meaning and discourse construction) (NACFLT, 2004).

Compared with TEM-4 (Test for English Majors Band 4 conducted at the end of English majors' fundamental stage-the second academic year) listening task, TEM-8 LSC poses a higher cognitive demand for test-takers who not only need to decode the acoustic input but also construct meaning to complete the task (Zou, 2004).

Insert Table 2

A TEM-8 qualified rater, who had marked the LSC task for more than three consecutive years, rated all the test-takers' answer sheets, according to the official TEM-8 scoring criteria. The maximum score for the LSC task was ten points. Content-irrelevant responses would result in deduction of one whole point while minor mistakes such as spelling errors would result in deduction of a half point unless the misspelled words are considered irrelevant or illegible.

Besides the TEM-8 LSC tasks, other research instruments include collecting think-aloud protocols based on TEM-8 LSC tasks and the participants' background questionnaires.

Data Collection Procedure

Both concurrent and immediate retrospective verbal reports are considered valid means to obtain individuals' thinking processes (Brown & Rodgers, 2002; Ericsson & Simon, 1993; Faerch & Kasper, 1983). In order to investigate what cognitive processes test-takers would experience while completing the LSC task, we used concurrent thinking-aloud as the main data collection tool. Below we describe the data collection procedure.

The sixteen participants were given a background questionnaire to fill in concerning their personal information such as age, gender, grade, and their perception of the LSC task of TEM-8. Thereafter, they were given instructions on what they needed to do during the main data collection. They were then given an opportunity to practice think-aloud using a practice version

of the LSC task of TEM-8. After a short break, the main data collection started. The tasks were administered as under TEM-8 exam conditions.

The think-aloud protocols were collected individually. First, a participant listened to the mini-lecture and simultaneously took notes, which lasted about 8 minutes. Second, s/he would be fully engaged in thinking aloud while tackling the LSC task after the mini-lecture. They were encouraged to verbalize "the heeded thoughts controlled by the central processor" (Ericsson and Simon, 1993). If there appeared a pause longer than ten seconds, a brief and clear probing was given, e.g. "Would you please go on?" It is important to note here that all the participants were only required to think aloud right on the spot rather than offering further interpretation of their thoughts or any comment on their performance under the exam condition or any assumption of a good test-taker's strategy. All the participants chose to use Chinese to verbalize their thoughts. As thinking-aloud would inevitably increase the time that the participants would normally spend in completing the LSC tasks, we allowed the participants to have additional time to ensure the production of the best quality think-aloud protocols under examination pressure. On average, the participants spent about 17 minutes (i.e. seven minutes more than the official ten minutes allowed in the TEM-8 tests) in completing their think-aloud process.

Finally, a brief semi-structured follow-up interview was conducted immediately after an individual participant's TAPs were collected. The interview lasted 10-15 minutes with each participant, aiming to seek clarifications on the participant's TAPs based on the field notes we took during TAP collection, with questions such as "What do you mean by saying this?" and "Can you confirm it please?". Figure 1 depicts the whole data collection procedure.

Insert Figure 1

Data Analysis

In this study, the coding scheme of all the cognitive processes of TEM-8 LSC was developed mainly based on Field's cognitive model of listening comprehension (2013). On the whole, the categories listed in Table 3 were a combination of the taxonomies of listening comprehension proposed by Field (2013) and the new ones that emerged from our TAP data.

The coding scheme covers decoding and selective attention, meaning and discourse construction, monitoring, decision-making, and test-wiseness strategies independent of the normal test-taking cognitive processes (see Table 3). The underlined processes in the left column of Table 3 were mainly based on Field's model (2013) and all the sub-processes were direct coding from the participants' TAP data.

Insert table 3

All the TAPs were coded twice in Nvivo 10.0 within five weeks. From week 1 to 2, the researcher finished the first coding of all TAPs; from week 3 to 4 an interval was inserted; in week 5, the second coding was completed. The intra-coder reliability coefficient of the two times of coding is 0.846. A language tester familiar with TEM-8 LSC task was invited to code 30% of the data for the second-time coding and the inter-coder reliability of that set of data was 0.895. Both intra-coder and inter-coder reliability levels were satisfactory.

Results

We report our findings from two perspectives: (1) an overview of test-takers' cognitive processes, and (2) their higher-level cognitive processes involved in the LSC tasks to explore how the different foci of the gaps impacted on test-takers' cognitive processes.

Participants' Cognitive Processes

Table 4 shows that all the 16 participants were engaged in decoding and selective attention, meaning and discourse construction and decision-making. Only two of them did not report anything related to the monitoring process. In terms of frequency, decoding and selective attention ranked first and meaning and discourse construction followed, while monitoring was the least mentioned.

Insert table 4

The overall reported cognitive processes of the 16 participants (see Table 5) do not present an observable difference except the one who only reported 7 cognitive processes. On the whole, quantitative nodes and references indicate similarities across different individuals. After decoding the input of the mini-lecture, including memorizing words from the recording (87 references, ranging from 0 to13 by individual), reading notes (141 references, ranging from 2 to 17), the test-takers' attention is rather focused on making sense of the notes, i.e. inferring from notes (16 references, ranging from 0 to 4) and contextualizing their notes in the gaps, including sub-processes such as inferring from the task (27 references, ranging from 0 to 5), interpreting (34 references, ranging from 0 to 6), building coherence (11 references, ranging from 0 to 4) and using schema (17 references, ranging from 0 to 3) in order to complete the comprehension task. Meanwhile, they are also constructing the discourse representation of the mini-lecture since the cloze task is given in an outline format that highly condenses the content of the lecture (14 references of summarization). In this case, test-takers would simultaneously be more inclined to make sense of the given outline while searching for suitable words in their notes to fill in the blanks.

Insert table 5

As an example, Table 6 showed Participant No. 2 reported in her TAPs that after decoding the input of the mini-lecture, her attention was rather focused on making sense of the notes she took and contextualizing them in the LSC task. Afterwards, she monitored her comprehension to double check whether her answer was both grammatically and semantically acceptable. Finally, she had to make a decision and fill in the blank with the words she believed were appropriate.

Insert table 6

Figure 2 shows the overall cognitive processes participants experienced in tackling TEM-8 LSC task. The main cognitive processes in Figure 2 include decoding and selective attention, meaning and discourse construction, monitoring process and decision-making represented by their detailed sub-processes which were listed in the order of references. We must note that decoding and selection process occupied 50.2% of the participants' TAP references and they unanimously employed sub-processes such as reading notes (16 sources, 141 references), memorizing words (14 sources, 87 references), reading the task (13 sources, 59 references) and parsing the task (13 sources, 35 references). As participants were completing the task after hearing the mini-lecture, information decoding and selection was derived mainly from notes (141 references) which suggested that they relied heavily on notes taken during the mini-lecture in order to tackle the TEM-8 LSC task. Meanwhile, participants had to decode the task while filling in the gaps, which added cognitive load.

Meaning and discourse construction is typical of higher-level cognitive processes. Participants made inferences from both notes and the task prompts (43 references) to build contextual coherence. For example, a participant reported: "So, reading the task and inferring from the following information, I can say getting assistance is one of the, eh, characteristics of

active learners…" Participants also reported how they interpreted new information in their TAPs (13 sources, 34 references). For example, a participant said "I think the word powerful is in contrast to less common expression" in order to interpret the meaning of the new information "powerful". Despite the low frequency of synthesizing (7 sources, 14 references), this sub-process was typical of discourse construction, through which participants must summarize the relevant details and clarify the main idea of the discourse. Participants' TAPs also revealed the importance of using schemata of world knowledge and the topic (10 sources, 17 references) that facilitated the cognitive processes such as meaning and discourse construction and monitoring. For example, a participant mentioned that she had learned educational theories and she knew clearly that the word "accommodate" meant "fit the new to the old", so her topic knowledge helped her with the answer.

In addition, two-way arrows between the cognitive processes imply 'recursions' of cognitive processing to illustrate the complicated path many participants used to arrive at final decisions. For example, in the process of meaning and discourse construction, participants would claim "I hear … in the mini-lecture" as a recursion to the decoding process in order to validate their customized meaning construction.

Insert figure 2

Interaction between Gap Types and Cognitive Processes and Targeting of Higher Cognitive Processes

In this section, we report the extent to which the four different gap types in the LSC tasks required or stimulated different levels and types of cognitive processing and how the participants responded differently to the four gap types (see Table 2). The extracts reported below are

examples that typically suggest the most dynamic interactions between gap types and the corresponding cognitive processes.

Gap Type 1: Key-point gaps. The following TAPs show how participants responded to the second gap of LSC that targeted one of the key-points of the lecture in the 2013 TEM-8 paper (see Appendix B for details).

Task extract.

(2) _____ and critical in thinking

i.e. information processing, e.g.

-connections between the known and the new information

(Key for reference: reflective)

TAP extracts.

Participant No. 6 *(Wrong answer; TEM-8 LSC score: 2.5)*: Then, the second one, critical in thinking. Um. I might fill in "quick". I wrote "quick" in my notes. And because there is "critical" here. It should also be an adjective.

Participant No. 1 *(Correct answer; TEM-8 LSC score: 3.5):* For the second blank, I heard "thinking critically, and reflect". So, I think it is, critical means thinking. I don't understand why the format here is "critical in thinking". Um, but since here is reading with purpose, I think I should fill in "reflecting". But indeed I can't understand why it is "critical in thinking". I don't know what it means. What I heard is "thinking critically".

Participant No. 7 *(Correct answer; TEM-8 LSC score: 4.5)*: The second, I heard "reflection and critical in thinking". Oh, reflective, critically, um...Actually, I heard both "reflection and reflective", but the second word here is critical, it is an

adjective. So I should use an adjective in the first place. Um. But when I read the task, I found "being able", and the previous one is "reading". Ok, I think I should add "being" here, it's "being reflective". Later, there's also "being able to get assistance", so I'd like to add "being" before "reflective".

Levels of processing. Participant No. 1 and 7 both said the word "heard" and it seems to suggest they tried to decode and search what they heard from the mini-lecture and they all parsed the structure. For example, Participant No. 1 compared the structure she heard "thinking critically" with the structure "critical in thinking" in the LSC task prompt and felt confused about the modification of the structure. This gap type has involved lower-level of processing: decoding, selective attention and parsing.

Different performances. The only difference between TAPs resulting in correct and wrong answers is that Participant No. 6 failed to select the key word "reflective" in the notes, which is sufficient to explain why she didn't make this answer right (See Appendix B for details).

Gap Type 2: Detail gaps. The following TAPs present how participants responded to the seventh gap which was designed to assess the participants' understanding of the detailed information of the lecture in the 2010 TEM-8 paper (see Appendix A for details).

Task extract.

1. proximity: physical distance between speakers

- closeness: intimacy or threat

- (7)_____: formality or absence of interest

(Key for reference: distance)

TAP extracts.

Participant No. 16 *(Wrong answer; TEM-8 LSC score: 1.5)*: … "threat" is before the seventh blank, so I need to find corresponding words after "threat" in my notes, I wrote down "lack of interest" and here's "absence of interest" in the outline, so I shall locate the answer and it should be the word before "lack of interest" in my notes, but I only wrote "dis", "dis" what, I didn't write a complete word. I cannot recall what this word is, so I shall leave it blank. I am not sure what this "dis" refers to.

Participant No. 13 *(Correct answer; TEM-8 LSC score: 4)*: Number seven. In my notes, I wrote "distance—formality", so the word in the gap should be "distance".

Participant No. 10 *(Correct answer; TEM-8 LSC score: 5)*: The seventh one is "keeping a distance". It is a contrast to closeness, and then the original wording is "if there's a difference, it means formality or absence, lack of interest of something…" The previous sentence is "closeness means intimacy or threat", and then the latter sentence is "if there's a difference, distance between two speakers, it means two speakers something..."

Participant No. 9 *(Correct answer; TEM-8 LSC score: 6)*: Number seven is "distance", because this one is in contrast to closeness in meaning and then, I also wrote the word down in notes.

Levels of processing. In all of the protocols, participants either mentioned they heard the word or wrote down the word in their notes. It is also clear that in order to fill details in the gaps, participants need to locate the answer in their notes, such as "it should be the word before 'lack of interest' in my notes", "I wrote distance—formality". Decoding process is reached across all these four protocols.

Different performances. Participant No. 16 said "but I only wrote 'dis'" and then could not make out the complete word; hence unsuccessful decoding of the word really posed difficulty in the meaning-construction process later. Participants No. 10 and No. 9 showed signs of semantic processing, as these two participants both mentioned: "It is a contrast (antonym) to closeness", which might offer a hint that higher cognitive processing is realized as they have attempted to build semantic coherence between phrases beyond simply noticing the sequence of them.

Gap Type 3: Inference gaps. The following TAPs present how participants responded to the seventh gap in the 2013 paper which was designed to assess the participants' ability in making inferences.

Task extract.

V. being able to question information

A. question what they read or hear

B. evaluate and (7) _____

(Key for reference: differ in conclusion/accommodate/adapt)

TAP extracts.

Participant No. 4 *(Wrong answer; TEM-8 LSC score: 3.5)*: Um, the seventh one, being able to question information. Question what they read or hear, evaluate and, um, and, I wrote here a word, that is "conclusion", after "evaluate", so, I think it might be "evaluate and conclude".

Participant No. 1 *(Wrong answer; TEM-8 LSC score: 3.5)*: Being able to question information. Question what they read or hear. Evaluate, evaluate and (pause 5 seconds) this, evaluate and...make...I heard here "when the new information does not fit in". Then,

so I think here, maybe, it should be "making, make connections". "Make connections". I think the meaning here should be a contrast between what they heard before and what they hear now.

Participant No. 2 *(Correct answer; TEM-8 LSC score: 4)*: I wrote "differ" but I am not sure. Yeah. Because judging from my notes, a comparatively suitable word in my notes should be "differ". I wrote down "question", I mean I wrote down the big point "question information", the fourth characteristic, but the basic structure of the notes of the lecture is always active learner blah-blah-blah and passive learner blah-blah-blah. So on the whole, I cannot spot very clear points, and I could only identify key words, when I see "differ" here as a key verb. I think this word could be the possible answer.

Participant No. 8 *(Correct answer; TEM-8 LSC score: 6.5)*: Being able to question information, evaluate and...Um. I will write "accommodate". Um. Also I don't write down the exact words, but I write "evaluate, read and hear", so I know it's about this part and then I wrote "the new failed to fit in what they know" and I know "fit the new to the old" is accommodation. I also write down "evaluate and read and hear", the new fails to fit in what they already knew.

Levels of processing. In all of the above protocols except No. 4's, participants have reached the level of meaning construction, which is evidenced by protocols such as "I think the meaning here should be…", "I see 'differ' here as a key verb", and "I know 'fit the new to the old' is accommodation". They all tried to figure out the exact meaning of the words to be filled in the gap.

Different performances. Participant No. 4 reported: "I wrote here a word, that is 'conclusion', after 'evaluate', so, I think it might be 'evaluate and conclude'", which clearly

shows the adjacency of words "evaluate and conclusion" in her notes; and in order to contextualize her notes, she changed the noun form "conclusion" into its verb form "conclude". However, there is no clear sign of meaning building in her protocol. Participant No. 1 interpreted the meaning of the word to be filled in the gap in her own way, attempting to build a logical relationship between the new and the old information, or as is put by her, "a contrast between what they heard before and what they hear now". The logic here makes sense except that the flow of thought is restricted due to the lack of relevant notes. Though Participants No. 2 and No. 8 got the correct answer, no evidence could indicate that No. 1 was inferior to them in terms of level of cognitive process. Participant No. 2 believed "differ" in her notes should be a key verb and this word could be the possible answer while she was contextualizing her own notes. However, she didn't clearly explain this part, which reveals that her meaning construction here was incomplete or unsuccessful. According to Participant No. 8's protocol, her content schema played a crucial role, as she later confirmed in the follow-up interview that she studied educational psychology, and hence she came up with the word "accommodate" without much effort.

On the whole, the analysis above has highlighted the importance of identifying key words, selecting and writing them down in notes for retrieval, building links between detailed propositional units and links between the input and the schema.

Gap Type 4: Synthesis gaps. The following TAPs present how participants responded to the eighth gap in the 2010 TEM-8 paper which was designed to assess the participants' summarization ability.

Task extract.

Proximity is person-, culture- and (8) _____-specific.

(Key for reference: situation)

TAP extracts.

Participant No. 12 *(Wrong answer; TEM-8 LSC score: 4)*: Number eight, person, culture specific, in this part, I remembered I heard the word "place", I'd write it here.

Participant No. 16 *(Wrong answer; TEM-8 LSC score: 1.5)*: Then the eighth one, I also follow the same sequence, and try to match the word with my notes, "in person, culture, and specific". For this one, because I wrote the word "informal" in my notes, and "proximity", I only wrote these words, and also the word "party", I decide to write down, I am not sure if it's correct, I write, I randomly guess it's "party" in this blank.

Participant No. 10 *(Correct answer; TEM-8 LSC score: 5)*: The eighth one is "situation". I wrote it down. Different culture, the first mentioned thing is "personal style", "culture-bounded", and then the following example is the importance of situation, so these three elements should be parallel to each other and I also highlighted them in my notes with asterisks. The speaker said something alike once again: "I want to emphasize the importance of what", so I here highlighted the key information with asterisks. And judging from the previous information, I believe these three should be parallel to each other and I'd put "situation" in the blank.

Level of processing. The gap summarizes the nature of proximity, aiming to reach participants' higher cognitive processing of discourse construction. Evidence from Participant No. 10's protocol suggests that the participant not only identified the key features pertaining to the nature of proximity, but also their semantic relationship to each other. Protocols of participant No. 12 and No. 16 indicate both of them merely reached decoding process and failed to get the right answer.

Different performances. A point worth noticing here is that No. 12 claimed he heard "place" while No. 16 mentioned "party". Neither of them explained why these words were suitable. Actually, the words they filled in the gaps are relevant to the context. However, two different "situations" could happen in the same "place", and "party" is only one of the hyponyms of "situation". Therefore, No. 12 and No. 16 didn't fully comprehend the accurate coherence between proximity, person, culture, situation and party, and they failed to reach the higher level of cognitive processing. On the other hand, Participant No. 10 exhibited a full range of discourse construction process from summarizing the concrete example to integrating the summary of the examples into the existent discourse structure.

According to the follow-up interview after TAPs were collected, we have also found that the LSC task can pose a challenge to participants' cognitive load in two aspects. Firstly, the participants found it particularly taxing to comprehend and take notes simultaneously. Even high achievers felt it challenging to comprehend complex sentences in the listening input that contain multiple idea units while they had to simultaneously take notes because of the high expectation of the usefulness of notes to complete the gaps at the second stage of the tasks. Secondly, the participants found the synthesis gaps the most challenging because they had to draw on and synthesize information from a number of sentences that were not necessarily presented in a linear or temporal order in the mini-lectures.

The above-listed TAPs have suggested that gap types in TEM-8 LSC tasks such as inference gaps and synthesis gaps managed to target the intended higher level of cognitive processes, namely meaning building and discourse construction defined by Field (2013: 100-103) and even some detail gaps intended for lower processes occasionally reached participants'

higher level of cognitive processing such as meaning building (refer to the section of *Detail gaps*).

Discussion and Conclusion

The main cognitive processes that the 16 participants underwent while completing the LSC tasks included decoding and selective attention, meaning and discourse construction, monitoring and decision-making. Amongst those processes, decoding and selective attention are considered as lower level of cognitive processes while the rest are believed to be higher level of cognitive processes. Decoding and selective attention as lower level of cognitive processes are represented by reading notes, memorizing words, reading the task, with highest references from participants in their TAPs, while meaning and discourse construction as higher level of cognitive processes have a number of sub-processes including interpreting, inferring, and synthesizing. Compared with participants' references of lower-level processes, not all of them have reached the higher-level processes or not all the intended higher-level processes are accessed by the task. The Cognitive Process Model of TEM-8 LSC (Figure 2) visualizes the cognitive processes participants have employed to tackle the task. Generally speaking, participants would resort heavily to lower-level processes in completion of the LSC task while higher-level processes are indispensable if they actively involve themselves in the interaction with the task and in making sense of it in order to achieve a high score. In the process of making sense of the task, participants usually automatically resort to their background knowledge of the world and topic. Whether the theme in the mini-lecture is relevant to their existing schema might influence the participants' ways of note-taking and their performance thereafter. The role of background knowledge needs to be considered by test designers and developers. However, there are also critical moments which suggested some participants might have reached higher level of

processes at the very beginning of the task. Therefore, participants' cognitive processes involved in TEM-8 LSC cannot be viewed as linear; rather it is iterative and multi-layered. Finally, the fact that most participants have employed the monitoring process suggests that they have actively engaged in this higher-level cognitive processing to ensure their best performance under the testing condition. Under examination conditions, participants need to make sure that their answers are both grammatically and semantically acceptable and therefore monitoring process is heavily loaded. Although our study was based on Field's model (2013) of cognitive processes of listening comprehension, our finding on the prominence of monitoring process in the LSC task can offer further insights into cognitive processes of listening comprehension under examination conditions.

Lower level cognitive processing was available in detail gaps and key-point gaps while higher level cognitive processing was evident in inference gaps and synthesis gaps. The higher frequency of lower level cognitive processes found in the current study somewhat echoes what Field argued in his 2009 study on the cognitive validity of lecture-based test tasks in the IELTS listening paper that much of the cognitive processing in the test condition was not deep enough. However, listening is different from reading as listeners do not have a chance to refer back to the original text, and hence higher frequency of lower cognitive processes is understandable. The overall quality of the LSC tasks seems to be affected by the extent to which inference and synthesis gaps can elicit participants' higher cognitive processes. The current study suggests interaction between cognitive processes and gap types in the LSC tasks is dynamic. How this interaction is realized varies between individual participants. As is shown in the data, the different types of gaps in the LSC tasks tapped into different sub-constructs of listening comprehension, and we therefore argue that any validation efforts on the LSC tasks need to

scrutinize not only the overall performance of test-takers for all the gaps but also their cognitive processing for completing each individual type of gaps. Furthermore, it is imperative to examine the relationships or interactions between gap types and cognitive processes and the extent to which the targeted cognitive processes are evident in students' test-taking processes.

Our TAP data suggested that the four gap types of TEM-8 LSC tasks have partially accessed the different levels of cognitive processes that test-designers intended to reach. This finding is similar to Shang's (2005) which concluded that test items requiring different levels of cognitive processing yielded correspondingly different levels of performance. In the current study, inference gaps have activated some participants' meaning construction by enabling them to not only decode but also contextualize new information. Synthesis gaps have partially achieved the intended higher-level cognitive process—meaning and discourse construction, by making participants synthesize a group of sentences of the original academic lecture, with their background knowledge fully engaged. The data also suggested that although the use of certain test-taking strategies might be helpful, those strategies alone were unlikely to help the participants attain correct answers.

The findings of the study are useful for test designers to better understand the underlying construct of LSC tasks and the challenges that English language learners face in comprehending academic lectures under examination conditions. We argue that successful comprehension of academic lectures requires identification of important points (see also Deroey, 2015; Flowerdew, 1994; Jung, 2003; Lynch, 2004; Tyler, 1992) and that assessment of test-takers' comprehension of important points should form an essential part of any academic listening comprehension test. From a test design perspective, we also have to consider a number of factors that may affect test-takers' performances on synthesis gaps, for example:

- the sequence of idea units that are central to the generation of the "synthesis,"
- the extent to which the idea units of the academic lecture provide sufficient clues for test-takers to generate the "synthesis,"
- whether the task prompts (in this case, the outline of the academic lecture) are available to test-takers when they listen to the lectures,
- the extent to which the task prompts provide clues for test-takers to infer/generate the "synthesis,"
- whether and to what extent test-takers can lift certain key words directly from the source lecture to successfully complete the tasks.

Meanwhile, we would like to acknowledge the various limitations of the study and point to the direction for future studies. Firstly, given the nature of the study that used think-aloud as the main data collection tool, the sample size was small and the participants were from one university only. They were not diverse enough in terms of their gender, academic backgrounds and English language proficiency (although their performance in the LSC tasks did vary to a large extent, see Table 1). We used two TEM-8 LSC tasks (2010 and 2013); the potential task-effects might be inevitable, although the two tasks were assumed to be equivalent in terms of their difficulty level according to the test provider – the TEM Committee. With a larger sample size of participants and more and different types of LSC tasks, we would be in a stronger position to generalize our findings and to better understand the relationships between participants' test performances and their engagement in different types of cognitive processing. Apart from the issues of sample size in relation to the use of think-aloud as the main data collection method, there are other limitations of think-aloud. It is inevitable that think-aloud could alter the test-taking process. Eye-tracking with think-aloud or eye-tracking alone could

provide critical information on how test-takers would use task prompts and integrate and summarize various sources in order to complete the different gap types in the LSC tasks. We aim to conduct our next project using eye-tracking technology to look into test-takers' cognitive processing of LSC tasks.

References

Anderson, A. & Lynch T. (1988). *Listening*. London: Oxford University.

Anderson, J. R. (1995). *Cognitive Psychology and Its Implications.* New York: Freeman.

Benson, M. J. (1994). Lecture listening in an ethnographic perspective. In J. Flowerdew (Ed.), *Academic listening: Research perspectives* (pp. 181-198). Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139524612.015

Brown, G. (2008). Selective listening. *System: An International Journal of Educational Technology and Applied Linguistics*, 36(1), 10-21. doi:10.1016/j.system.2007.11.002

Brown, G. and Yule, G. (1983). *Discourse Analysis*. Cambridge: Cambridge University Press.

Brown, J. D. & Rodgers. T. S. (2002). *Doing Second Language Research*. Oxford: Oxford University Press.

Buck, G. (1991). The testing of listening comprehension: an introspective study. *Language Testing*, 8(1), 67-91. doi: 10.1177/026553229100800105

Buck, G. (1992). Listening comprehension: Construct validity and trait characteristics. *Language Learning*, 42(3), 313-357. doi: 10.1111/j.1467-1770.1992.tb01339.x

Buck, G. (2001). *Assessing Listening*. Cambridge: Cambridge University Press.

Carroll, J. B. (1972). Defining language comprehension: Some speculations. In R. O. Freedle & J. B. Carroll (Eds.), *Language comprehension and the acquisition of knowledge* (pp.1-29). Washington, DC: Winston.

Clark, H. H., and Clark, E.V. (1977). *Psychology and Language: An Introduction to Psycholinguistics*. New York: Harcourt Brace Jovanovich.

Coniam, D. (1993). Co-text or no-text? A study of an adapted cloze technique for testing comprehension/summary skills. *Hong Kong Papers in Linguistics and Language Teaching*, 16(1), 1-10.

Courchêne, R., & Bayliss, D. (1995). Summary cloze: What is it? What does it measure? In R. Courchêne, S. Burger, C. Cornaire, R. LeBlanc, S. Paribakht & H. Seguin (Eds.), *Twenty-five years of second language teaching at the University of Ottawa* (pp. 305-327). Ottawa: Second Language Institute, University of Ottawa.

Cumming, A., Laia, C. & Cho, H. (2016). Students' writing from sources for academic purposes: A synthesis of recent research. *Journal of English for Academic Purposes*, 23(12), 47-58. doi: 10.1016/j.jeap.2016.06.002

Deroey, K. L. B. (2015). Marking importance in lectures: Interactive and textual orientation. *Applied Linguistics*, 36(1), 51–72. doi: 10.1093/applin/amt029

Ericsson, K. A., & Simon, H. A. (1993). *Protocol Analysis: Verbal Reports as Data* (2nd Ed.). Cambridge, MA: MIT Press.

Faerch, G. & Kasper, G. (1983). *Strategies in Interlanguage Communication*. London: Longman.

Field, J. (2009). A cognitive validation of the lecture-listening component of the IELTS listening paper. In P. Thompson & R. Badger & M. Carey & S. Devi & J. Field et al. (Eds.),

International English Language Testing System: volume 9: Research reports 2009. British Council/IELTS Australia/ Cambridge ESOL.

Field, J. (2011). Into the mind of the academic listener. *Journal of English for Academic Purposes*, 10(2), 102-112. doi: 10.1016/j.jeap.2011.04.002

Field, J. (2013). Cognitive validity. In A. Geranpayeh & L. Taylor (Eds.), *Examining listening* (pp. 77-151). Cambridge: Cambridge University Press.

Flowerdew, J. (1994). Research of relevance to second language lecture comprehension-an overview. In J. Flowerdew (Ed.), *Academic listening: Research perspectives* (pp. 7-29). Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139524612.004

Gellert, A. & Elbro, C. (2013). Cloze tests may be quick, but are they dirty? Development and preliminary validation of a cloze test of reading comprehension. *Journal of Psychoeducational Assessment*, 31(1), 16 –28. doi: 10.1177/0734282912451971

Goh, C. (2005). Second language listening expertise. In K. Johnson (Ed.), *Expertise in second language learning and teaching* (pp. 64–84). Basingstoke, UK: Palgrave Macmillan. doi: 10.1057/9780230523470_4

Goh, C. (2010). Listening as process: Learning activities for self-appraisal and self-regulation. In N. Harwood (Ed.), *Materials in ELT: Theory and practice* (pp. 179–206). Cambridge: Cambridge University Press.

Greene, B. B. (2001). Testing reading comprehension of theoretical discourse with Cloze. *Journal of Research in Reading*, 24(1), 82-98. doi: 10.1111/1467-9817.00134

Hansen, C. & Jensen, C. (1994). Evaluating lecture comprehension. In Flowerdew, J. (Ed.). *Academic listening: Research perspectives* (pp. 241-268). Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139524612.019

Hulstijn, J. (2010). Basic language cognition and the construct of language proficiency: theoretical and empirical perspectives. Paper presented at the 32nd Language Testing Research Colloquium, (LTRC). Cambridge, April 2010.

Hyon, S. (1997). *Models of Lecture Discourse: Applications for Academic Listening and Future Research Directions. Colloquium on Academic Listening within the EAP Curriculum*. Distributed by ERIC Clearinghouse, [Washington D.C.]

Jin, Y. & Fan, J. (2011). Test for English Majors (TEM) in China. *Language Testing*, 28(4), 589-596. doi: 10.1177/0265532211414852

Jung, E. H. (2003). The role of discourse signaling cues in second language listening comprehension. *Modern Language Journal*, 87(4), 562-577. doi: 10.1111/1540-4781.00208

Kiely, (2004). Learning to critique in EAP. Journal of English for Academic Purposes, 3(3), 211-227. doi: 10.1016/j.jeap.2003.11.004

Kintsch, W., & Kozminsky, E. (1977). Summarizing stories after reading and listening. *Journal of Educational Psychology*, 69(5), 491-499. doi: 10.1037/0022-0663.69.5.491

Kintsch, W. & van Dijk, T. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5), 363-394. doi: 10.1037/0033-295X.85.5.363

Lee, J. J. (2009). Size matters: An exploratory comparison of small- and large-class university lecture introductions. *English for Specific Purposes*, 28(1), 42-57. doi: 10.1016/j.esp.2008.11.001

Lewkowicz, J. A. (1992). Testing listening comprehension using listening summary cloze. *JALT Journal*, 14(1), 7-22.

Lynch, T. (2004). *Study Listening: A Course in Listening to Lectures and Note-Taking*. Cambridge: Cambridge University Press.

Lynch, T. (2006). Acadmic listening: marrying top and bottom. In E. Usó-Juan and A. Martinez-Flor (Eds.), *Current trends in the development and teaching of the four Language skills* (pp. 91-110). Mouton de Gruyter: Berlin.

Lynch, T. (2011). Academic listening in the 21st century: Reviewing a decade of research. *Journal of English for Academic Purposes*, 10(2), 79-88. doi: 10.1016/j.jeap.2011.03.001

NACFLT (2004). *Syllabus for TEM8*. Shanghai: Shanghai Foreign Language Education Press.

Pashler, H. & Johnson, J. C. (1998). Attentional limitations in dual task performance. In Pashler, H. (Ed.) *Attention* (pp. 155-189). Hove: Psychology Press.

Richards, J. C. (1983). Listening comprehension: approach, design, procedure. *TESOL Quarterly*, 17(2), 219-240. doi: 10.2307/3586651

Rost, M. (1990). *Listening in Language Learning*. London: Longman.

Rost, M. (1994). Online summaries as representations of lecture understanding. In Flowerdew, J. (Ed.) *Academic listening: Research perspectives* (pp.93-127). Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139524612.009

Rowley-Jolivet, E. (2002). Visual discourse in scientific conference papers: A genre-based study. *English for Specific Purposes*, 21(1), 19-40. doi: 10.1016/S0889-4906(00)00024-7

Shang, H. (2005). An investigation of cognitive operations on L2 Listening Comprehension Performance: An Exploratory study. *International Journal of Listening*, 19(1): 51-62. doi: 10.1080/10904018.2005.10499074

Storey, P. (1997). Examining the test-taking process: a cognitive perspective on the discourse cloze test. *Language Testing*, 14(2), 214-231. doi: 10.1177/026553229701400205

Taylor, L. & Geranpayeh A. (2011). Assessing listening for academic purposes: Defining and operationalizing the test construct. *Journal of English for Academic Purposes*, 10(2), 89-101. doi: 10.1016/j.jeap.2011.03.002

Taylor, L. (2013). *Testing Reading through Summary: Investigating Summary Completion Tasks for Assessing Reading Comprehension Ability*. Cambridge: Cambridge University Press.

Tyler, A. (1992). Discourse structuring and the perception of incoherence in international teaching assistants' spoken discourse. *TESOL Quarterly*, 26(4), 713–729. doi: 10.2307/3586870

Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching*, 40(3), 191-210. doi: 10.1017/S0261444807004338

Waggoner, M. (1984). The new technologies versus the lecture tradition in higher education: Is change possible? *Educational Technology*, 24(3), 7-12.

Young, L. (1994). University lectures—macro-structure and micro-features. In J. Flowerdew (Ed.), *Academic listening: Research perspectives* (pp. 159-176). Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139524612.013

Yu, G. (2005). *Towards a model of using summarization tasks as a measure of reading comprehension*. Bristol, Unpublished PhD Dissertation.

Yu, G. (2013). The use of summarization tasks: some lexical and conceptual analyses. *Language Assessment Quarterly*, 10(1), 96-109. DOI: 10.1080/15434303.2012.750659

Zou, S. (2004). How to achieve interactiveness in listening tests-Designing the new TEM-8 listening subtest. *Computer-assisted Foreign Language Education*, 12(6), 33-37. doi:10.3969/j.issn.1001-5795.2004.06.007

Appendices

Appendix A. TEM-8 LSC (2010)

Instructions:

Complete the gap-filling task. Some of the gaps below may require a maximum of THREE words. Make sure the word(s) you fill in is (are) both grammatically & semantically acceptable. You may refer to your notes.

Paralinguistic Features of Language

In face-to-face communication speakers often alter their tones of voice or change their physical postures in order to convey messages. These means are called paralinguistic features of language, which fall into two categories.

I. First category: vocal paralinguistic features

A. (1)_____: to express attitude or intention (1) _____

B. examples

1. whispering: need for secrecy
2. breathiness: deep emotion
3. (2)_____: unimportance (2) _____
4. nasality: anxiety
5. extra lip rounding greater intimacy

II. Second category: physical paralinguistic features

A. facial expressions

1. (3)_____ (3) _____

- smiling: signal of pleasure or welcome

2. less common expressions

- eyebrow raising: surprise or interest
- lip biting: (4)_____ (4)_____

B. gesture

- Gestures are related to culture.

1. British culture

- shrugging shoulders: (5)_____ (5) _____
- scratching head: puzzlement

2. other cultures

- placing hand upon heart: (6)_____ (6)_____
- pointing at nose: secret

C. proximity, posture and echoing

1. proximity: physical distance between speakers

- closeness: intimacy or threat
- (7)_____: formality or absence of interest (7)_____

Proximity is person-, culture- and (8)_____-specific. (8)_____

2. Posture

- hunched shoulders or a hanging head: to indicate (9) _____ (9)_____
- direct level eye contact: to express an open or challenging attitude

3. echoing

- definition: imitation of similar posture
- (10)_____: aid in communication (10)_____

-conscious imitation: mockery

Test-takers' Cognitive Processes during a Listen-to-Summarize Cloze Task

Keys to TEM-8 LSC (2010)

1. function/purpose/intention
2. huskiness
3. common expressions
4. thought/uncertainty/thinking
5. indifference/lack of knowledge
6. honesty
7. distance
8. situation
9. mood
10. unconscious imitation

Script of TEM-8 LSC (2010)

Paralinguistic Features of Language

Good morning everyone. Today, we'll continue our discussion on describing language. Last week we examined such features of, of language as grammar, vocabulary, the sounds of language, etc. In this lecture, we'll look at another important aspect of language. Perhaps some of you may wonder what is this important aspect of language. Let me tell you, it refers to features of communication that takes place without the use of grammar or vocabulary. They are called paralinguistic features of language. These features fall into two broad categories, those that involve the voice and those that involve the body.

Now, the first category is what we call vocal paralinguistic features. Vocal features are actually tones of voice. While they are, perhaps, not central to meaning in communication in the same way as grammar or vocabulary, they may nevertheless convey attitude or intention in some way. Let me give you some examples. The first is whispering, which indicates the need for secrecy. The second is breathiness; this is to show deep emotion. The third is huskiness, which is to show unimportance. The fourth is nasality; this, um, is to indicate anxiety. The last is extra lip rounding, which expresses greater intimacy, especially with babies, for example. So, we can see

that there are a number of ways of altering our tone of voice, and when we do this consciously, we do it to create different effects in communication.

Now, let's come to the second category: physical paralinguistic features, which involves the body. In addition to conveying meanings with tone of voice, we can also express our intention through the ways in which we use our bodies. You may ask: what are the ways then? Let me cite some brief examples. The expression on our face, the gestures we make, and even proximity or the way we sit are some of the ways we send powerful messages about how we feel or what we mean.

Let me explain some of these in more detail.

First, facial expression. Facial expression is a powerful conveyor of meaning. We all know smiling is an almost universal signal of pleasure or welcome. But there are other facial expressions that may not be so common. For instance, raising eyebrows suggests that you are surprised or interested in something. Other facial actions such as biting your lip, which indicates that you are deep in thinking or are uncertain about something, compressing the lips, which shows that you are making decisions, and a visible clenching of the teeth to show that you are angry, are all powerful conveyors of meaning, too.

The second in this category is gesture. You see, we use gesture to indicate a wide range of meanings, though I have to emphasize that the actual gestures we use may be specific to particular cultures. That is to say, different cultures have their own favored gestures in conveying meaning. Here, a few examples may show you how powerful gestures can be. In British English behavior, shrugging shoulders may indicate an attitude of *I don't care* or *I don't know*. Crossing your arms may indicate relaxation, but it can also powerfully show you are bored. Waving can mean welcome and farewell, whereas scratching your head may indicate that you're at a loss. In

other cultures, placing your hand upon your heart is to indicate that you are telling the truth. Pointing your finger at your nose means: it's a secret. That's why we say that gestures are culture-bound.

The third is proximity, posture and echoing. Proximity refers to the physical distance between speakers. This can indicate a number of things and can also be used to consciously send messages about intent. Closeness, for example, indicates intimacy or threat to many speakers. But distance may show formality or lack of interest. Once again, I'd like to say, proximity is also both a matter of personal style and is often culture-bound. So, what may seem normal to a speaker from one culture may appear unnecessarily close or distant to a speaker from another. And standing close to someone may be quite appropriate in some situations such as an informal party, but completely out of place in other situations, such as a meeting with a superior.

Next, posture. Posture means the way in which someone holds his or her body, especially the back, shoulders and head when standing, walking or sitting. A few examples. Hunched shoulders and a hanging head give a powerful indication of whether the person is happy or not. A lowered head when speaking to a superior (with or without eye contact) can convey the appropriate relationship in some cultures. On the other hand, direct level eye contact changes the nature of the interaction and can be seen as either open or challenging.

Last, echoing. Now, what is echoing? Let me start with an example. Some of you may have noticed this phenomenon in your experience. When two people are keen to agree with each other, they would likely, though unconsciously, adopt the same posture as if in imitation of each other. They sit or stand in the same manner. When used in this way, echoing appears to complement the verbal communication. Of course, when such imitation is carried out consciously, it often indicates that someone is mocking at another speaker.

Test-takers' Cognitive Processes during a Listen-to-Summarize Cloze Task

Ok. In today's lecture we looked at some paralinguistic features such as tone of voice, gesture and posture. These features, together with linguistic features of language like grammar or vocabulary, are all part of the way we communicate with each other in face-to-face encounters. In our next lecture, we'll watch some video material, and see how people actually use paralinguistic means in communication, to express their intention or desire or mood.

Appendix B. TEM-8 LSC (2013)

Instructions:

Complete the gap-filling task. Some of the gaps below may require a maximum of THREE words. Make sure the word(s) you fill in is (are) both grammatically & semantically acceptable. You may refer to your notes.

What Do Active Learners Do

There are differences between active learning and passive learning.
Characteristics of active learners:

I. reading with purpose
  A. before reading: setting goals
  B. while reading: (1) _____ (1) _____
II. (2) _____ and critical in thinking (2) _____
  i.e. information processing, e.g.
  -connections between the known and the new information
  -identification of (3) _____ concepts (3) _____
  -judgment on the value of (4) _____ (4) _____
III. active in listening
  A. ways of note-taking: (5) _____ (5) _____
  B. before note-taking: listening and thinking
IV. being able to get assistance
  A. reason 1: knowing comprehension problems because of (6) _____ (6) _____
  B. reason 2: being able to predict study difficulties
V. being able to question information
  A. question what they read or hear
  B. evaluate and (7) _____ (7) _____
VI. last characteristic
  A. attitude toward responsibility
    -active learners: accept
    -passive learners: (8) _____ (8) _____
  B. attitude toward (9) _____ (9) _____
    -active learners: evaluate and change behavior
    -passive learners: no change in approach

Relationship between skill and will: will is more important in (10) _____ (10)_____
Lack of will leads to difficulty in college learning.

Keys to TEM-8 LSC (2013)

1. checking (their) understanding/comprehension
2. reflective/reflecting (on information)
3. unfamiliar/unknown
4. reading materials/what you read
5. comprehensive and organized/structured
6. monitoring/understanding
7. differ in conclusion/adapt/change/alter the conclusion
8. blame others
9. poor/unsatisfactory/bad performance
10. active learning/college study/being active learners

Test-takers' Cognitive Processes during a Listen-to-Summarize Cloze Task

Script of TEM-8 LSC (2013)

What Do Active Learners Do

Good morning. Today I'll discuss what is active learning and, uh…what do active learners do? In order to define active learning, I'll look at the differences between active learning and passive learning by examining six characteristics of active learners and contrasting them with those of passive learners. Ok, let's start.

The first characteristic: active learners tend to read with the purpose of understanding and remembering. I bet that no one deliberately sits down to read with the purpose of not understanding the text. However, I am sure that some of you have been in a situation, probably more than once, where you read, quote and unquote, "an assignment," closed the text and thought, "what the world was that about?" When you interact with the text in this manner, you are reading passively. Active readers, on the other hand, set goals before they read, and check their understanding as they read. When they finish, they can explain the main points, and know that they have understood what they have read.

Now, the second characteristic of active learners is to reflect on information and think critically. Being reflective is an important part of active learning, because it means that you are thinking about the information, in other words, you are processing the information. For instance, you may make connections between the new information and what you already know, identify concepts that you may not understand very well, or evaluate the importance of what you are reading. An active learner reflects constantly in this way. In contrast, passive learners may read the text and listen to lectures, and even understand most of what is read and heard, but they do not take that crucial next step of actually thinking about it.

Let's move on to the third one. The third characteristic is to listen actively by taking comprehensive notes in an organized way, like what you should be doing now. We lecturers are always amazed at the number of students who are engaged in activities other than listening and note-taking in their lecture classes. We've seen students reading newspapers, doing an assignment for another class or chatting with the classmates. Perhaps the all time winner for passive learning, however, was a student who regularly came to my class with a pillow and fell asleep. Unlike these students, active learners are engaged learners. They listen actively to the professor for the entire class period and they write down as much information as possible. To be an active note-taker, you must be more than simply present. You have to think about the information before you write.

The fourth characteristic is to get assistance when they are experiencing problems. Because active learners are constantly monitoring their understanding, they know when their comprehension breaks down, and they ask for help before they become lost. In addition, active learners often predict the courses, or even particular concepts within courses that may give them trouble. They have a plan in mind for letting assistance should they need it. Active learners may seek assistance from their professors, or peers. Although passive learners may seek help at some point, it is often too little, too late. In addition, because passive learners do not reflect and think critically, they often don't even realize that they need help.

The next characteristic is to question information; this means that active learners raise questions on information that they read and hear, while passive learners accept both the printed page and the words of their professors as truth. Of course, active learners don't question everything, but they do evaluate what they read and hear. When new information fails to fit in

with what they already know, they may differ in the conclusions they draw or in the inferences they make.

The last characteristic, which I think is the most fundamental one, is to accept much of the responsibility for learning. Active learners understand that the responsibility for learning must come from within, while passive learners often want to blame others for their lack of motivation, poor performance, time management problems, and other difficulties that they might experience. When active learners don't perform as well as they had hoped, they evaluate why they didn't do well, and change those studying behaviors the next time. Passive learners on the other hand, often approach every course in the same manner and then get angry with professors when their performance is poor. It is only when students accept the responsibility for their own learning that they can truly be called "active learners".

So, from what I've said so far, you can see that being an active learner involves both skill and will. By skill, I mean the tools to handle the studying and learning demands placed on you, like how to read with purpose, when and where to get assistance if you are having difficulty. By will, I mean the desire and motivation to follow through. Here I'd like to emphasize that skill is nothing without will. For example, you may have a friend who is knowledgeable, but not motivated in the classroom. Even though he reads widely and can intelligently discuss a variety of issues, he does little schoolwork and rarely studies. In other words, students such as these may have the skills to do well, but for some reason they simply do not have the will. And because skill and will go hand in hand, unmotivated students, those who do not have the will may experience difficulty in college.

Okay, today we've discussed the differences between an active learner and a passive one, and some useful study strategies that may eventually help you become an active learner.